Secure Identification and QKD in the Bounded-Quantum-Storage Model*

Ivan Damgård[1], Serge Fehr[2,**], Louis Salvail[3], and Christian Schaffner[2,***]

[1] DAIMI, Aarhus University, Denmark
ivan@cs.au.dk
[2] Centrum Wiskunde & Informatica (CWI) Amsterdam, The Netherlands
{s.fehr|c.schaffner}@cwi.nl
[3] Université de Montréal (DIRO), QC, Canada
salvail@iro.umontreal.ca

Abstract. We consider the problem of secure identification: user U proves to server S that he knows an agreed (possibly low-entropy) password w, while giving away as little information on w as possible, namely the adversary can exclude at most one possible password for each execution of the scheme. We propose a solution in the bounded-quantum-storage model, where U and S may exchange qubits, and a dishonest party is assumed to have limited quantum memory. No other restriction is posed upon the adversary. An improved version of the proposed identification scheme is also secure against a man-in-the-middle attack, but requires U and S to additionally share a high-entropy key k. However, security is still guaranteed if one party loses k to the attacker but notices the loss. In both versions of the scheme, the honest participants need no quantum memory, and noise and imperfect quantum sources can be tolerated. The schemes compose sequentially, and w and k can securely be re-used. A small modification to the identification scheme results in a quantum-key-distribution (QKD) scheme, secure in the bounded-quantum-storage model, with the same re-usability properties of the keys, and without assuming authenticated channels. This is in sharp contrast to known QKD schemes (with unbounded adversary) without authenticated channels, where authentication keys must be updated, and unsuccessful executions can cause the parties to run out of keys.

1 Introduction

Secure Identification. Consider two parties, a *user* U and a *server* S, who share a common secret-key (or password or Personal Identification Number PIN) w. In order to obtain some service from S, U needs to convince S that he is the legitimate user U by "proving" that he knows w. In practice—think of how you prove to the ATM that you know your PIN—such a proof is often done simply by announcing w to S. This indeed guarantees that a dishonest user U^* who does not know w cannot identify himself as U, but of course incurs the risk that U might reveal w to a malicious server S^* who may now impersonate U. Thus, from a secure identification scheme we also require that a dishonest server S^* obtains (essentially) no information on w.

There exist various approaches to obtain secure identification schemes, depending on the setting and the exact security requirements. For instance zero-knowledge proofs (and some weaker versions), as initiated by Feige, Fiat and Shamir [FS86, FFS87], allow for secure identification. In a more sophisticated model, where we allow the common key w to be of low entropy and additionally consider a man-in-the-middle attack, we can use techniques from password-based key-agreement

* A preliminary version of this paper appeared in *Advances in Cryptology—CRYPTO 2007* [DFSS07].
** Supported by the Dutch Organization for Scientific Research (NWO).
*** supported by EU fifth framework project QAP IST 015848 and the NWO VICI project 2004-2009.

(like [KOY01, GL03]) to obtain secure identification schemes. Common to these approaches is that security relies on the assumption that some computational problem (like factoring or computing discrete logs) is hard and that the attacker has limited computing power.

Our Contribution. In this work, we take a new approach: we consider quantum communication, and we develop two identification schemes which are information-theoretically secure under the *sole* assumption that the attacker can only reliably store quantum states of limited size. This model was first considered in [DFSS05]. On the other hand, the honest participants only need to send qubits and measure them immediately upon arrival, no quantum storage or quantum computation is required. Furthermore, our identification schemes are robust to both noisy quantum channels and imperfect quantum sources. Our schemes can therefore be implemented in practice using existing technology.

The first scheme is secure against dishonest users and servers but not against a man-in-the-middle attack. It allows the common secret-key w to be non-uniform and of low entropy, like a human-memorizable password. Only a user knowing w can succeed in convincing the server. In any execution of this scheme, a dishonest user or server cannot learn more on w than excluding one possibility, which is unavoidable. This is sometimes referred to as *password-based* identification. The second scheme requires in addition to w a uniformly distributed high-entropy common secret-key k, but is additionally secure against a man-in-the-middle attack. Furthermore, security against a dishonest user or server holds as for the first scheme even if the dishonest party knows k (but not w). This implies that k can for instance be stored on a smart card, and security of the scheme is still guaranteed even if the smart card gets stolen, assuming that the affected party notices the theft and thus does not engage in the scheme anymore. Both schemes compose sequentially, and w (and k) may be safely re-used super-polynomially many times, even if the identification fails (due to an attack, or due to a technical failure).

A small modification of the second identification scheme results in a quantum-key-distribution (QKD) scheme secure against bounded-quantum-memory adversaries. The advantage of the proposed new QKD scheme is that no authenticated channel is needed and the attacker can *not* force the parties to run out of authentication keys. The honest parties merely need to share a password w and a high-entropy secret-key k, which they can safely re-use (super-polynomially many times), independent of whether QKD succeeds or fails. Furthermore, like for the identification scheme, losing k does not compromise security as long as the loss is noticed by the corresponding party. One may think of this as a quantum version of password-based authenticated key exchange. The properties of our solution are in sharp contrast to all known QKD schemes without authenticated channels (which do not pose any restrictions on the attacker). In these schemes, an attacker can force parties to run out of authentication keys by making the QKD execution fail (e.g. by blocking some messages). Worse, even if the QKD execution fails only due to technical problems, the parties can still run out of authentication keys after a short while, since they cannot exclude that an eavesdropper was in fact present. This problem is an important drawback of QKD implementations, especially of those susceptible to single (or few) point(s) of failure [EPT03].

Other Approaches. We briefly discuss how our identification schemes compare with other approaches. We have already given some indication on how to construct *computationally* secure identification schemes. This approach typically allows for very practical schemes, but requires some unproven complexity assumption. Another interesting difference between the two approaches: whereas for (known) computationally-secure password-based identification schemes the underlying

computational hardness assumption needs to hold indefinitely, the restriction on the attacker's quantum memory in our approach only needs to hold *during* the execution of the identification scheme, actually only at one single point during the execution. In other words, having a super-quantum-storage-device at home in the basement only helps you cheat at the ATM if you can communicate with it on-line quantumly – in contrast to a computational solution, where an off-line super-computer in the basement can make a crucial difference.

Furthermore, obtaining a satisfactory identification scheme requires *some* restriction on the adversary, even in the quantum setting: considering only passive attacks, Lo [Lo97] showed that for an unrestricted adversary, no password-based quantum identification scheme exists. In fact, Lo's impossibility result only applies if the user U is guaranteed not to learn anything about the outcome of the identification procedure. We can argue, however, that a different impossibility result holds even without Lo's restriction: We first show that secure computation of a classical AND gate (in which both players learn the output) can be reduced to a password-based identification scheme. The reduction works as follows. Let w_0, w_0' and w_1 be three distinct elements from $\mathcal{W}$. If Alice has private input $x_A = 0$ then she sets $w_A = w_0$ and if $x_A = 1$ then she sets $w_A = w_1$, and if Bob has private input $x_B = 0$ then he sets $w_B = w_0'$ and if $x_B = 1$ then he sets $w_B = w_1$. Then, Alice and Bob run the identification scheme on inputs w_A and w_B, and if the identification is rejected, the output is set to 0 while if it is accepted, the output is set to 1. Security of the identification scheme is easily seen to imply security of the AND computation. Now, the secure computation of an AND gate—with statistical security and using quantum communication—can be shown to require a superpolynomial number of rounds if the adversary is unbounded [NPS07]. Therefore, the same must hold for a secure password-based identification scheme.[1]. In fact, in very recent work [BCS09], using the definitions from [FS09], it is shown that the whole password of the honest player leaks to the dishonest player.

Another alternative approach is the classical bounded-storage model [Mau90, CCM98, ADR02]. In contrast to our approach, only classical communication is used, and it is assumed that the attacker's *classical* memory is bounded. Unlike in the quantum case where we do not need to require the honest players to have any quantum memory, the classical bounded-storage model requires honest parties to have a certain amount of memory which is related to the allowed memory size of the adversary: if two legitimate users need n bits of memory in an identification protocol meeting our security criterion, then an adversary must be bounded in memory to $O(n^2)$ bits. The reason is that given a secure password-based identification scheme, one can construct (in a black-box manner) a key-distribution scheme that produces a one-bit key on which the adversary has an (average) entropy of $\frac{1}{2}$. On the other hand it is known that in any key-distribution scheme which requires n bits of memory for legitimate players, an adversary with memory $\Omega(n^2)$ can obtain the key except for an arbitrarily small amount of remaining entropy [DM04]. It follows that password-based identification schemes in the classical bounded-storage model can only be secure against adversaries with memory at most $O(n^2)$. This holds even for identification schemes with only passive security and without security against man-in-the-middle attacks. Roughly, the reduction works as follows. Alice and Bob agree on a public set of two keys $\{w_0, w_1\}$. Alice picks $a \in_R \{0,1\}$, Bob picks $b \in_R \{0,1\}$, and they run the identification scheme with keys w_a and w_b respectively.

[1] In fact, we believe that the proof from [NPS07] can be extended to cover secure computation of equality of strings, which is equivalent to password-based identification. This would mean that we could prove the impossibility result directly, without the detour via a secure AND computation.

The outcome of the identification is then made public from which Bob determines a. We argue that if the identification fails, i.e. $a \neq b$, then a is a secure bit. Thus, on average, a has entropy (close to) $\frac{1}{2}$ from an eavesdropper's point of view. Consider $w' \notin \{w_0, w_1\}$. By the security property of the identification scheme, Alice and thus also a passive eavesdropper Eve cannot distinguish between Bob having used w_b or w'. Similarly, we can then switch Alice's key w_a to w_{1-a} and Bob's switched key w' to w_{1-b} without changing Eve's view. Thus, Eve cannot distinguish an execution with $a = 0$ from one with $a = 1$ if $a \neq b$.

This limitation of the classical bounded-storage model is in sharp contrast with what we achieve in this paper, the honest players need no quantum memory at all while our identification scheme remains secure against adversaries with quantum memory linear in the total number of qubits sent. The same separation between the two models was shown for OT and bit commitment [DFSS05, DFR$^+$07].

Finally, if one settles for the bounded-quantum-storage model, then in principle one could take a generic construction for general two-party secure-function-evaluation (SFE) based on OT together with the OT scheme from [DFSS05, DFR$^+$07] in order to implement a SFE for string equality and thus password-based identification. However, this approach leads to a highly impractical solution, as the generic construction requires many executions of OT, whereas our solution is comparable with *one* execution of the OT scheme from [DFSS05, DFR$^+$07]. Furthermore, SFE does not automatically take care of a man-in-the-middle attack, thus additional work would need to be done using this approach.

Subsequent Work. The difficulty of storing quantum information can also be modeled differently from assuming a bound on the physical number of qubits an adversary can control. In the more realistic noisy-quantum-storage model put forward in [WST08], all incoming qubits can be stored by an adversary but are subject to storage noise. Assuming a simple storage strategy, one can show that the protocols in the current paper remain secure [STW08], whereas it is unknown if security still holds in case of more sophisticated storage strategies [KWW09].

If the storage limitation on the adversary fails to hold, it is easy to see that not only will our security proofs fail, but in fact the protocol we propose can be broken quite efficiently. However, it was recently shown, in [DFL$^+$09], how to add a "preamble" to the protocol using a commitment scheme based on a computational assumption. It is shown in [DFL$^+$09] that to break the resulting protocol, an adversary must have both large quantum memory *and* large computing power.

2 Preliminaries

2.1 Notation and Terminology

Quantum States. We assume the reader's familiarity with basic notation and concepts of quantum information processing [NC00]. In this paper, the computational or $+$-basis is defined by the pair $\{|0\rangle, |1\rangle\}$ (also written as $\{|0\rangle_+, |1\rangle_+\}$). The pair $\{|0\rangle_\times, |1\rangle_\times\}$ denotes the diagonal or $\times$-basis, where $|0\rangle_\times = (|0\rangle + |1\rangle)/\sqrt{2}$ and $|1\rangle_\times = (|0\rangle - |1\rangle)/\sqrt{2}$. We write $|x\rangle_\theta = |x_1\rangle_{\theta_1} \otimes \cdots \otimes |x_n\rangle_{\theta_n}$ for the n-qubit state where string $x = (x_1, \ldots, x_n) \in \{0,1\}^n$ is encoded in bases $\theta = (\theta_1, \ldots, \theta_n) \in \{+, \times\}^n$.

The behavior of a (mixed) quantum state in a register E is fully described by its density matrix ρ_E. In order to simplify language, we tend to be a bit sloppy and use E as well as ρ_E as "naming" for the quantum state. We often consider cases where a quantum state E may depend on

some classical random variable X (from a finite set $\mathcal{X}$) in that the state is described by the density matrix ρ_E^x if and only if $X = x$. For an observer who has only access to the state E but not to X, the behavior of the state is determined by the density matrix $\rho_E := \sum_x P_X(x)\rho_E^x$, whereas the joint state, consisting of the classical X and the quantum state E, is described by the density matrix $\rho_{XE} := \sum_x P_X(x)|x\rangle\langle x| \otimes \rho_E^x$, where we understand $\{|x\rangle\}_{x\in\mathcal{X}}$ to be the standard (orthonormal) basis of $\mathbb{C}^{|\mathcal{X}|}$. More general, for any event $\mathcal{E}$ (defined by $P_{\mathcal{E}|X}(x) = P[\mathcal{E}|X=x]$ for all x), we write

$$\rho_{XE|\mathcal{E}} := \sum_x P_{X|\mathcal{E}}(x)|x\rangle\langle x| \otimes \rho_E^x \quad \text{and} \quad \rho_{E|\mathcal{E}} := \mathrm{tr}_X(\rho_{XE|\mathcal{E}}) = \sum_x P_{X|\mathcal{E}}(x)\rho_E^x \ . \tag{1}$$

We also write $\rho_X := \sum_x P_X(x)|x\rangle\langle x|$ for the quantum representation of the classical random variable X (and similarly for $\rho_{X|\mathcal{E}}$). This notation extends naturally to quantum states that depend on several classical random variables, defining the density matrices ρ_{XYE}, $\rho_{XYE|\mathcal{E}}$, $\rho_{YE|X=x}$ etc. We tend to slightly abuse notation and write $\rho_{YE}^x = \rho_{YE|X=x}$ and $\rho_{YE|\mathcal{E}}^x = \rho_{YE|X=x,\mathcal{E}}$, as well as $\rho_E^x = \mathrm{tr}_Y(\rho_{YE}^x)$ and $\rho_{E|\mathcal{E}}^x = \mathrm{tr}_Y(\rho_{YE|\mathcal{E}}^x)$.[2] Note that writing $\rho_{XE} = \mathrm{tr}_Y(\rho_{XYE})$ and $\rho_E = \mathrm{tr}_{X,Y}(\rho_{XYE})$ is consistent with the above notation. We also write $\rho_{XE|\mathcal{E}} = \mathrm{tr}_Y(\rho_{XYE|\mathcal{E}})$ and $\rho_{E|\mathcal{E}} = \mathrm{tr}_{X,Y}(\rho_{XYE|\mathcal{E}})$, where one has to be aware that in contrast to (1), here the state E may depend on the event $\mathcal{E}$ (namely via Y), so that, e.g., $\rho_{E|\mathcal{E}} = \sum_x P_{X|\mathcal{E}}(x)\rho_{E|\mathcal{E}}^x$. Given a quantum state E that depends on a classical random variable X, by saying that there exists a random variable Y such that ρ_{XYE} satisfies some condition, we mean that ρ_{XE} can be understood as $\rho_{XE} = \mathrm{tr}_Y(\rho_{XYE})$ for some ρ_{XYE} (with classical Y) and that ρ_{XYE} satisfies the required condition.[3]

X is independent of E (in that ρ_E^x does not depend on x) if and only if $\rho_{XE} = \rho_X \otimes \rho_E$, which in particular implies that no information on X can be learned by observing only E. Similarly, X is random and independent of E if and only if $\rho_{XE} = \frac{1}{|\mathcal{X}|}\mathbb{I} \otimes \rho_E$, where $\frac{1}{|\mathcal{X}|}\mathbb{I}$ is the density matrix of the fully mixed state of suitable dimension. Finally, if two states like ρ_{XE} and $\rho_X \otimes \rho_E$ are ε-close in terms of their trace distance $\delta(\rho,\sigma) = \frac{1}{2}\mathrm{tr}(|\rho - \sigma|)$, which we write as $\rho_{XE} \approx_\varepsilon \rho_X \otimes \rho_E$, then the real system ρ_{XE} "behaves" as the ideal system $\rho_X \otimes \rho_E$ except with probability ε in that for any evolution of the system no observer can distinguish the real from the ideal one with advantage greater than ε [RK05]. As ε can be interpreted as an error probability, we typically require ε to be *negligible* in a security parameter n, denoted as $\varepsilon = negl(n)$. A security parameter is a natural number n given as input to all players in our protocols, and a probability is said to be negligible in n if for any polynomial p, it is smaller than $1/p(n)$ for all sufficiently large n.

Conditional Independence. We also need to express that a random variable X is (close to) independent of a quantum state E *when given a random variable* Y. This means that when given Y, the state E gives no (or little) additional information on X. Formally, this is expressed by requiring that ρ_{XYE} equals (or is close to) $\rho_{X\leftrightarrow Y\leftrightarrow E}$, which is defined as[4]

$$\rho_{X\leftrightarrow Y\leftrightarrow E} := \sum_{x,y} P_{XY}(x,y)|x\rangle\langle x| \otimes |y\rangle\langle y| \otimes \rho_E^y \, .$$

[2] The density matrix $\rho_{E|\mathcal{E}}^x$ describes the quantum state E in the case that the event $\mathcal{E}$ occurs and X takes on the value x. The corresponding holds for the other density matrices considered here.

[3] This is similar to the case of distributions of classical random variables where given X the existence of a certain Y is understood that there exists a certain joint distribution P_{XY} with $\sum_y P_{XY}(\cdot,y) = P_X$.

[4] The notation is inspired by the classical setting where the corresponding independence of X and Z given Y can be expressed by saying that $X \leftrightarrow Y \leftrightarrow Z$ forms a Markov chain.

In other words, $\rho_{XYE} = \rho_{X\leftrightarrow Y\leftrightarrow E}$ precisely if $\rho_E^{x,y} = \rho_E^y$ for all x and y. To further illustrate its meaning, notice that if the Y-register is measured and value y is obtained, then the state $\rho_{X\leftrightarrow Y\leftrightarrow E}$ collapses to $(\sum_x P_{X|Y}(x|y)|x\rangle\langle x|) \otimes \rho_E^y$, so that indeed no further information on x can be obtained from the E-register. This notation naturally extends to $\rho_{X\leftrightarrow Y\leftrightarrow E|\mathcal{E}}$ simply by considering $\rho_{XYE|\mathcal{E}}$ instead of ρ_{XYE}. Explicitly, $\rho_{X\leftrightarrow Y\leftrightarrow E|\mathcal{E}} = \sum_{x,y} P_{XY|\mathcal{E}}(x,y)|x\rangle\langle x| \otimes |y\rangle\langle y| \otimes \rho_{E|\mathcal{E}}^y$.

The notion of conditional independence has been introduced in [DFSS07] (a classical version was independently proposed in [CW08]) and used as a convenient tool in subsequent papers [FS09, BCS09]. In this paper we will use the following property of conditional independence whose proof is given in Appendix A.1.

Lemma 2.1. *For any event $\mathcal{E}$, the density matrix $\rho_{X\leftrightarrow Y\leftrightarrow E}$ can be decomposed into*

$$\rho_{X\leftrightarrow Y\leftrightarrow E} = P[\mathcal{E}]^2 \cdot \rho_{X\leftrightarrow Y\leftrightarrow E|\mathcal{E}} + (1 - P[\mathcal{E}]^2) \cdot \tau$$

for some density matrix τ. Furthermore, if $\mathcal{E}$ is independent of X and Y, then

$$\rho_{X\leftrightarrow Y\leftrightarrow E} = P[\mathcal{E}] \cdot \rho_{X\leftrightarrow Y\leftrightarrow E|\mathcal{E}} + P[\bar{\mathcal{E}}] \cdot \rho_{X\leftrightarrow Y\leftrightarrow E|\bar{\mathcal{E}}}\,.$$

(Conditional) Smooth Min-Entropy. Different notions of conditional (smooth) min-entropy have been proposed in the literature; we briefly specify here the variant that is convenient for us. Let X and Y be random variables, over respective finite alphabets $\mathcal{X}$ and $\mathcal{Y}$, with joint distribution P_{XY}. The *conditional min-entropy* of X given Y is defined as the negative logarithm of the guessing probability of X given Y: $H_{\min}(X|Y) := -\log\big(p_{\text{guess}}(X|Y)\big)$ where

$$p_{\text{guess}}(X|Y) := \sum_y P_Y(y) \max_x P_{X|Y}(x|y) = \sum_y \max_x P_{XY}(x,y)$$

and log denotes the binary logarithm (here and throughout the paper). More generally, we define $H_{\min}(X\mathcal{E}|Y)$ for any event $\mathcal{E}$ as $H_{\min}(X\mathcal{E}|Y) := -\log\big(p_{\text{guess}}(X\mathcal{E}|Y)\big)$ where[5]

$$p_{\text{guess}}(X\mathcal{E}|Y) := \sum_y P_Y(y) \max_x P_{X\mathcal{E}|Y}(x|y) = \sum_y \max_x P_{XY\mathcal{E}}(x,y)\,.$$

The *conditional smooth min-entropy* $H_{\min}^\varepsilon(X|Y)$ is then defined as

$$H_{\min}^\varepsilon(X|Y) := \max_{\mathcal{E}} H_{\min}(X\mathcal{E}|Y)$$

where the max is over all events $\mathcal{E}$ with $P[\mathcal{E}] \geq 1 - \varepsilon$.

Obviously, the unconditional versions of smooth and non-smooth min-entropy are obtained by using an "empty" Y; furthermore the above notions extend naturally to $H_{\min}(X|Y,\mathcal{E})$ and $H_{\min}^\varepsilon(X|Y,\mathcal{E})$ for any event $\mathcal{E}$ by considering the corresponding conditional joint distribution $P_{XY|\mathcal{E}}$.

[5] $p_{\text{guess}}(X\mathcal{E}|Y)$ can be understood as the optimal probability in guessing X *and* have $\mathcal{E}$ occur, when given Y.

2.2 Tools

MIN-ENTROPY-SPLITTING . A technical tool, which will come in handy, is the following entropy-splitting lemma, which may also be of independent interest. Informally, it says that if for a list of random variables, every pair has high (smooth) min-entropy, then all of the random variables except one must have high (smooth) min-entropy. The proof is given in Appendix A.2.

Lemma 2.2 (Entropy-Splitting Lemma). *Let $\varepsilon \geq 0$. Let $X_1, \ldots, X_m$ and Z be random variables such that $H_{\min}^{\varepsilon}(X_i X_j|Z) \geq \alpha$ for all $i \neq j$. Then there exists a random variable V over $\{1, \ldots, m\}$ such that for any* independent *random variable W over $\{1, \ldots, m\}$ with $H_{\min}(W) \geq 1$,*

$$H_{\min}^{2m\varepsilon}(X_W|VWZ, V \neq W) \geq \alpha/2 - \log(m) - 1\,.$$

QUANTUM UNCERTAINTY RELATION. At the very core of our security proofs lies (a special case of) the quantum uncertainty relation from [DFR+07][6], that lower bounds the (smooth) min-entropy of the outcome when measuring an arbitrary n-qubit state in a random basis $\theta \in \{0,1\}^n$.

Theorem 2.3 (Uncertainty Relation [DFR+07]). *Let E be an arbitrary fixed n-qubit state. Let Θ be uniformly distributed over $\{+,\times\}^n$ (independent of E), and let $X \in \{0,1\}^n$ be the random variable for the outcome of measuring E in basis Θ. Then, for any $\lambda > 0$, the conditional smooth min-entropy is lower bounded by*

$$H_{\min}^{\varepsilon}(X|\Theta) \geq \Big(\frac{1}{2} - 2\lambda\Big)n$$

with $\varepsilon \leq 2^{-\sigma(\lambda)n}$ and $\sigma(\lambda) = \frac{\lambda^2 \log(e)}{32(2-\log(\lambda))^2}$.

Thus, ignoring negligibly small "error probabilities" and linear fractions that can be chosen arbitrarily small, the outcome of measuring any n-qubit state in a random basis has $n/2$ bits of min-entropy, given the basis.

PRIVACY AMPLIFICATION. Finally, we recall the quantum-privacy-amplification theorem of Renner and König [RK05]. The version we use here follows immediately from [Ren05, Corollary 5.6.1] by applying the chain rule for min- and max-entropy [Ren05, Lemma 3.2.9] and using the equivalence, as shown in [KRS08], of the quantum and the classical notion of (smooth) conditional min-entropy. Recall that a class $\mathcal{F}$ of hash functions from $\mathcal{X}$ to $\mathcal{Y}$ is called (strongly) universal-2 if for any $x \neq x' \in \mathcal{X}$, and for F uniformly distributed over $\mathcal{F}$, the collision probability $P[F(x) = F(x')]$ is upper bounded by $1/|\mathcal{Y}|$, respectively, for the strong notion, the random variables $F(x)$ and $F(x')$ are uniformly and independently distributed over $\mathcal{Y}$.

Theorem 2.4. *Let X and Z be random variables distributed over $\mathcal{X}$ and $\mathcal{Z}$, respectively, and let E be a q-qubit state that may depend on X and Z. Let F be the random and independent choice of a member of a universal-2 class of hash functions $\mathcal{F}$ from $\mathcal{X}$ into $\{0,1\}^{\ell}$. Then, for any $\varepsilon > 0$*

$$\delta\big(\rho_{F(X)FZE}, \tfrac{1}{2^{\ell}}\mathbb{I} \otimes \rho_{FZE}\big) \leq \frac{1}{2}\, 2^{-\frac{1}{2}\left(H_{\min}^{\varepsilon}(X|Z) - q - \ell\right)} + 2\varepsilon\,.$$

[6] In [DFR+07], a stricter notion of conditional smooth min-entropy was used, which in particular implies the bound as stated here.

3 The Identification Scheme

3.1 The Setting

We assume that the honest user U and the honest server S share some key $w \in \mathcal{W}$ (which we think of as a password), where the choice of w is described by the random variable W. An identification protocol is now simply any protocol for U and S using classical and/or quantum communication where the parties are both given as input a security parameter n and (in the honest case) the password w, and where S outputs accept or reject in the end.

We do not require $\mathcal{W}$ to be very large (i.e. $|\mathcal{W}|$ does not have to be lower bounded by the security parameter in any way), and w does not necessarily have to be uniformly distributed in $\mathcal{W}$. So, we may think of w as a human-memorizable password or PIN code. The goal of this section is to construct an identification scheme that allows U to "prove" to S that he knows w. The scheme should have the following security properties: a dishonest server S^* learns essentially no information on w beyond that he can come up with a guess w' for w and learns whether $w' = w$ or not, and similarly a dishonest user succeeds in convincing the verifier essentially only if he guesses w correctly, and if his guess is incorrect then the only thing he learns is that his guess is incorrect. This in particular implies that as long as the entropy of W is large enough, the identification scheme may be safely repeated. Finally, it must of course be the case that S accepts the legitimate user who has the correct password. More formally, we require the following:

Definition 3.1. *An execution by honest U, S on input w for both parties results in S accepting, except with negligible probability (as a function of n).*

Definition 3.2. *We say that an identification protocol for two parties U, S is secure for the user with error ε against (dishonest) server S^* if the following is satisfied: whenever the initial state of S^* is independent of W, the joint state $\rho_{WE_{\mathsf{S}^*}}$ after the execution of the protocol is such that there exists a random variable W' that is independent of W and such that*

$$\rho_{WW'E_{\mathsf{S}^*}|W'\neq W} \approx_\varepsilon \rho_{W\leftrightarrow W'\leftrightarrow E_{\mathsf{S}^*}|W'\neq W}.$$

Definition 3.3. *We say that an identification protocol for two parties U, S is secure for the server with error ε against (dishonest) user U^* if the following is satisfied: whenever the initial state of a dishonest user U^* is independent of W, there exists W' (possibly $\perp$), independent of W, such that if $W \neq W'$ then S accepts with probability at most ε, and if $W = W'$ then S accepts with certainty. Furthermore, the common state $\rho_{WE_{\mathsf{U}^*}}$ after the execution of the protocol (including S's announcement to accept or reject) satisfies*

$$\rho_{WW'E_{\mathsf{U}^*}|W'\neq W} \approx_\varepsilon \rho_{W\leftrightarrow W'\leftrightarrow E_{\mathsf{U}^*}|W'\neq W} \quad .$$

If these definitions are satisfied for a small ε, we are guaranteed that whatever a dishonest party does is essentially as good as trying to guess W by some arbitrary (but independent) W' and learning whether the guess was correct or not, but nothing beyond that. Such a property is obviously the best one can hope for, since an attacker may always honestly execute the protocol with a guess for W and observe whether the protocol was successful.

We would like to point out that the above security definitions, and in fact any security claim in this paper, guarantees *sequential self-composability*, as the output state is guaranteed to have the

same independency property (for any fixed choice of W') as is required from the input state (except if the attacker guesses W). Moreover, it is shown in [FS08a, FS09] that our definitions imply a "real/ideal" world definition given in [FS09]. More specifically, it is shown that a protocol satisfying our information theoretic conditions implements a natural ideal identification functionality, and by the composition theorem from [FS09], this means that the protocol composes sequentially in a classical environment, i.e. the quantum protocol can be treated as the ideal functionality when analyzing a more complicated classical outer protocol.

It should be noted that security for user and server is usually not sufficient for application in practice of an identification protocol. A problem occurs if the honest user and server are interacting and an attacker can manipulate the communication, i.e., do a "man-in-the-middle" attack, and observe the reaction of the honest parties. This scenario is not covered by the above definitions, and indeed it turns out that the simplest version of our protocol is not secure against such an attack. Nevertheless, the problem can be solved and we address it in Section 4.

3.2 The Intuition

The scheme we propose is related to the (randomized) 1-2 OT scheme of [DFR$^+$07]. In that scheme, Alice sends $|x\rangle_\theta$ to Bob, for random $x \in \{0,1\}^n$ and $\theta \in \{+,\times\}^n$. Bob then measures everything in basis $+$ or $\times$, depending on his choice bit c, so that he essentially knows half of x (where Alice used the same basis as Bob) and has no information on the other half (where Alice used the other basis), though, at this point, he does not know yet which bits he knows and which ones he does not. Then, Alice sends θ and two hash functions to Bob, and outputs the hash values s_0 and s_1 of the two parts of x, whereas Bob outputs the hash value s_c that he is able to compute from the part of x he knows. It is proven in [DFR$^+$07] that no dishonest Alice can learn c, and for any quantum-memory-bounded dishonest Bob, at least one of the two strings s_0 and s_1 is random for Bob.

This scheme can be extended by giving Bob more options for measuring the quantum state. Instead of measuring all qubits in the $+$ or the $\times$ basis, he may measure using m different strings of bases, where any two possible basis-strings have large Hamming distance. Then Alice computes and outputs m hash values, one for each possible basis-string that Bob might have used. She reveals θ and the hash functions to Bob, so he can compute the hash value corresponding to the basis that he has used, and no other hash value. Intuitively, such an extended scheme leads to a randomized 1-m OT.

The scheme can now be transformed into a secure identification scheme as follows, where we assume (wlog) that $\mathcal{W} = \{1,\ldots,m\}$. The user U, acting as Alice, and the server S, acting as Bob, execute the randomized 1-m OT scheme where S "asks" for the string indexed by his key w, such that U obtains random strings $s_1,\ldots,s_m$ and S obtains s_w. Then, to do the actual identification, U sends s_w to S, who accepts if and only if it coincides with his string s_w. Intuitively, such a construction is secure against a dishonest server since unless he asks for the right string (by guessing w correctly) the string U sends him is random and thus gives no information on w. On the other hand, a dishonest user does not know which of the m strings S asked for and wants to see from him. We realize this intuitive idea in the next section. In the actual protocol, U does not have to explicitly compute all the s_i's, and also we only need a single hash function (to compute s_w). We also take care of some subtleties, for instance that the s_i are not necessarily random if Alice (i.e. the user) is dishonest.

3.3 The Basic Scheme

Let $\mathfrak{c} : \mathcal{W} \to \{+,\times\}^n$ be the encoding function of a binary code of length n with $m = |\mathcal{W}|$ codewords and minimal distance d. $\mathfrak{c}$ can be chosen such that n is linear in $\log(m)$ or larger, and d is linear in n. Furthermore, let $\mathcal{F}$ and $\mathcal{G}$ be strongly universal-2 classes of hash functions[7] from $\{0,1\}^n$ to $\{0,1\}^\ell$ and from $\mathcal{W}$ to $\{0,1\}^\ell$, respectively, for some parameter ℓ. For $x \in \{0,1\}^n$ and $I \subseteq \{1,\ldots,n\}$, we define $x|_I \in \{0,1\}^n$ to be the restriction of x to the coordinates x_i with $i \in I$. If $|I| < n$ then applying $f \in \mathcal{F}$ to $x|_I$ is to be understood as applying f to $x|_I$ padded with sufficiently many 0's.

Q-ID:
1. U picks $x \in_R \{0,1\}^n$ and $\theta \in_R \{+,\times\}^n$, and sends state $|x\rangle_\theta$ to S.
2. S measures $|x\rangle_\theta$ in basis $c = \mathfrak{c}(w)$. Let x' be the outcome.
3. U picks $f \in_R \mathcal{F}$ and sends θ and f to S. Both compute $I_w := \{i : \theta_i = \mathfrak{c}(w)_i\}$.
4. S picks $g \in_R \mathcal{G}$ and sends g to U.
5. U computes and sends $z := f(x|_{I_w}) \oplus g(w)$ to S.
6. S accepts if and only if $z = z'$ where $z' := f(x'|_{I_w}) \oplus g(w)$.

It is trivial that the protocol satisfies Definition 3.1. In addition, we have:

Proposition 3.4 (User security). *Assume that the size of the quantum memory of dishonest server* S^* *is at most* q *at step 3 of Q-ID, and that* $H_{\min}(W) \geq 1$. *Then Q-ID is secure for the user with error* ε *against* S^* *according to Definition 3.2, where*

$$\varepsilon = 2^{-\frac{1}{2}((\frac{1}{4}-\lambda)d-\log(m)-q-\ell-1)} + 2^{-(\sigma(\lambda)d-\log(m)-3)}$$

for an arbitrary $0 < \lambda < \frac{1}{4}$.

Note that $\sigma(\lambda)$ was defined earlier in the claim of the uncertainty relation. To understand what the result on ε means, note that using a family of asymptotically good codes, we can assume that d grows linearly with the main security parameter n, while still allowing m (the number of passwords) to be exponential in n. So we may choose the parameters such that $\frac{d}{n}, \frac{\log(m)}{n}, \frac{q}{n}$ and $\frac{\ell}{n}$ are all constants. The result above now says that ε is exponentially small as a function of n if these constants are chosen in such a way that for some $0 < \lambda < \frac{1}{4}$, it holds that $(\frac{1}{4}-\lambda)\frac{d}{n} - \frac{\log(m)}{n} - \frac{q}{n} - \frac{\ell}{n} > 0$ and $\sigma(\lambda)\frac{d}{n} - \frac{\log(m)}{n} > 0$. See Theorem 3.6 for a choice of parameters that also take server security into account. If we are willing to assume that $\log(m)$ is sublinear in n, which may be quite reasonable is case we use short passwords that humans can remember, the condition further simplifies to $\frac{d}{4n} - \frac{q}{n} - \frac{\ell}{n} > 0$.

Proof. We consider and analyze a *purified* version of *Q-ID* where in step 1, instead of sending $|x\rangle_\theta$ to S^* for a random x, U prepares a fully entangled state $2^{-n/2}\sum_x |x\rangle|x\rangle$ and sends the second register to S^* while keeping the first. Then, in step 3 when the memory bound has applied, he measures his register in the random basis $\theta \in_R \{+,\times\}^n$ in order to obtain x. Standard arguments imply that this purified version produces exactly the same common state, consisting of the classical information on U's side and S^*'s quantum state.

[7] Actually, we only need $\mathcal{G}$ to be *strongly* universal-2.

Recall that before step 3 is executed, the memory bound applies to S^*, meaning that S^* has to measure all but q of the qubits he holds, which consists of his initial state and his part of the EPR pairs. Before doing the measurement, he may append an ancilla register and apply an arbitrary unitary transform. As a result of S^*'s measurement, S^* gets some outcome y, and the common state collapses to a $(n+q)$-qubit state (which depends on y), where the first n qubits are with U and the remaining q with S^*. The following analysis is for a fixed y, and works no matter what y is.

We use upper case letters W, X, Θ, F, G and Z for the random variables that describe the respective values w, x, θ etc. in an execution of the purified version of *Q-ID*. We write $X_j = X|_{I_j}$ for any j, and we let E'_{S^*} be S^*'s q-qubit state at step 3, after the memory bound has applied. Note that W is independent of X, Θ, F, G and E'_{S^*}.

For $1 \leq i \neq j \leq m$, fix the value of X, and correspondingly of X_i and X_j, at the positions where $\mathfrak{c}(i)$ and $\mathfrak{c}(j)$ coincide, and focus on the remaining (at least) d positions. The uncertainty relation (Theorem 2.3) implies that the restriction of X to these positions has $(\frac{1}{2} - 2\lambda)d$ bits of ε'-smooth min-entropy given Θ, where $\varepsilon' \leq 2^{-\sigma(\lambda)d}$ and $0 < \lambda < \frac{1}{2}$ arbitrary. Since every bit in the restricted X appears in one of X_i and X_j, the pair X_i, X_j also has $(\frac{1}{2} - 2\lambda)d$ bits of ε'-smooth min-entropy given Θ. The Entropy Splitting Lemma 2.2 implies that there exists W' (called V in Lemma 2.2) such that if $W \neq W'$ then X_W has $(\frac{1}{4} - \lambda)d - \log(m) - 1$ bits of $2m\varepsilon'$-smooth min-entropy given W and W' (and Θ). Privacy amplification then guarantees that $F(X_W)$ is ε''-close to random and independent of F, W, W', Θ and E'_{S^*}, conditioned on $W \neq W'$, where $\varepsilon'' = \frac{1}{2} \cdot 2^{-\frac{1}{2}(d/4 - \lambda d - \log(m) - 1 - q - \ell)} + 4m\varepsilon'$. It follows that $Z = F(X_W) \oplus G(W)$ is ε''-close to random and independent of F, G, W, W', Θ and E'_{S^*}, conditioned on $W \neq W'$.

Formally, we want to upper bound $\delta(\rho_{WW'E_{\mathsf{S}^*}|W' \neq W},\ \rho_{W \leftrightarrow W' \leftrightarrow E_{\mathsf{S}^*}|W' \neq W})$. Since the output state E_{S^*} is, without loss of generality, obtained by applying some unitary transform to the set of registers $(Z, F, G, W', \Theta, E'_{\mathsf{S}^*})$, the distance above is equal to the distance between $\rho_{WW'(Z,F,G,\Theta,E'_{\mathsf{S}^*})|W' \neq W}$ and $\rho_{W \leftrightarrow W' \leftrightarrow (Z,F,G,\Theta,E'_{\mathsf{S}^*})|W' \neq W}$. We then get:

$$\begin{aligned} &\rho_{WW'(Z,F,G,\Theta,E'_{\mathsf{S}^*})|W' \neq W} \approx_{\varepsilon''} \tfrac{1}{2^\ell}\mathbb{I} \otimes \rho_{WW'(F,G,\Theta,E'_{\mathsf{S}^*})|W' \neq W} \\ &= \tfrac{1}{2^\ell}\mathbb{I} \otimes \rho_{W \leftrightarrow W' \leftrightarrow (F,G,\Theta,E'_{\mathsf{S}^*})|W' \neq W} \approx_{\varepsilon''} \rho_{W \leftrightarrow W' \leftrightarrow (Z,F,G,\Theta,E'_{\mathsf{S}^*})|W' \neq W} \ , \end{aligned}$$

where approximations follow from privacy amplification and the exact equality comes from the independency of W, which, when conditioned on $W' \neq W$, translates to independency given W'. The claim follows with $\varepsilon = 2\varepsilon''$. □

Proposition 3.5 (Server security). *If $H_{\min}(W) \geq 1$, then Q-ID is secure for the server with error ε against any U^* according to Definition 3.3, where $\varepsilon = m^2/2^\ell$.*

The formal proof is given below. The idea is the following. We let U^* execute *Q-ID* with a server that is *unbounded* in quantum memory. Such a server can obviously obtain x and thus compute $s_j = f(x|_{I_j}) \oplus g(j)$ for all j. Note that s_w is the message z that U^* is required to send in the last step. Now, if the s_j's are all distinct, then z uniquely defines w' such that $z = s_{w'}$, and thus S accepts if and only if $w' = w$, and U^* does not learn anything beyond. The strong universal-2 property of g guarantees that the s_j's are all distinct except with probability $m^2/2^\ell$.

Proof. Again, we consider a slightly modified version. We let U^* interact with a server that has *unbounded* quantum memory and does the following. Instead of measuring $|x\rangle_\theta$ in step 2 in basis

c, it stores the state and measures it after step 3 in basis θ (and obtains x). This modified version produces the same common state $\rho_{WE_{\mathsf{U}^*}}$ as the original scheme, since the only difference between the two is when and in what basis the qubits at positions $i \notin I_w$ are measured, which does not effect the execution in any way.

We use the upper case letters W, X, Θ, F, G and Z for the random variables that describe the respective values w, x, θ etc. in an execution of the modified version of *Q-ID*. Furthermore, we define $S_j := F(X|_{I_j}) \oplus G(j)$ for $j = 1, \ldots, m$. Note that $Z' = S_W$ represents the value z' used by S in the last step. Let $\mathcal{E}$ be the event that all S_j's are distinct. By the strong universal-2 property, and since G is independent of X and F, the S_j's are pairwise independent and thus it follows from the union bound that $\mathcal{E}$ occurs except with probability at most $m(m-1)/2 \cdot 1/2^\ell \leq m^2/2^{\ell+1}$.

Let E'_{U^*} be U^*'s quantum state after the execution of *Q-ID* but *before* he learns S's decision to accept or reject. We may assume that the values of all random variables X, Θ, F, G, Z and the S_j's are known/given to U^*, i.e., we consider them as part of E'_{U^*}. Furthermore, we may assume that Z is one of the S_j's, i.e. that $Z = S_{W'}$ for a random variable W'. Indeed, if $Z \neq S_j$ for all j then we set $W' := \perp$ and S's decision is "reject", no matter what W is, and U^* obviously learns no information on W at all. By the way we have defined W', is clear that S accepts if $W = W'$.

Note that E'_{U^*} is independent of W by assumption on U^*'s initial state (in Definition 3.3) and by definition of the random variables X, Θ etc. Since $\mathcal{E}$ is determined by the S_j's (which are part of E'_{U^*}), this holds also when conditioning on $\mathcal{E}$. This then translates to the independence of E'_{U^*} from W when given W', conditioned on $W' \neq W$ and $\mathcal{E}$.

We now consider U^*'s state E_{U^*} *after* he has learned S's decision. If $W' \neq W$ and all S_j's are distinct then S rejects with probability 1. Hence, conditioned on the events $W' \neq W$ and $\mathcal{E}$, U^*'s state E_{U^*} remains independent of W given W'. Define $p := P[\mathcal{E}|W' \neq W]$ and $\bar{p} := P[\bar{\mathcal{E}}|W' \neq W] = 1 - p$, where $\bar{\mathcal{E}}$ is the complementary event to $\mathcal{E}$. Recall that $P[\bar{\mathcal{E}}] \leq m^2/2^{\ell+1}$, and therefore $\bar{p} \leq P[\bar{\mathcal{E}}]/(1 - P[W' = W]) \leq 2P[\bar{\mathcal{E}}] \leq m^2/2^\ell$, where the second-last inequality follows from the independence of W and W', and from the condition on $H_{\min}(W)$. Note that $\bar{p}$ upper bounds the probability that S accepts in case $W' \neq W$, proving the first claim. From the above it follows that

$$\begin{aligned}\rho_{WW'E_{\mathsf{U}^*}|W' \neq W} &= p \cdot \rho_{WW'E_{\mathsf{U}^*}|\mathcal{E}, W' \neq W} + \bar{p} \cdot \rho_{WW'E_{\mathsf{U}^*}|\bar{\mathcal{E}}, W' \neq W} \\ &= p \cdot \rho_{W \leftrightarrow W' \leftrightarrow E_{\mathsf{U}^*}|\mathcal{E}, W' \neq W} + \bar{p} \cdot \rho_{WW'E_{\mathsf{U}^*}|\bar{\mathcal{E}}, W' \neq W} \,.\end{aligned}$$

Furthermore, it is not too hard to see that $\mathcal{E}$ is independent of W and W', and thus also when conditioned on $W' \neq W$. Lemma 2.1 hence implies that

$$\rho_{W \leftrightarrow W' \leftrightarrow E_{\mathsf{U}^*}|W' \neq W} = p \cdot \rho_{W \leftrightarrow W' \leftrightarrow E_{\mathsf{U}^*}|\mathcal{E}, W' \neq W} + \bar{p} \cdot \rho_{W \leftrightarrow W' \leftrightarrow E_{\mathsf{U}^*}|\bar{\mathcal{E}}, W' \neq W} \,.$$

By definition of the metric $\delta(\cdot, \cdot)$, and because it cannot be bigger than 1, the distance between the two states is at most $\bar{p} \leq m^2/2^\ell$. □

We call an identification scheme *ε-secure against impersonation attacks* if the protocol is secure for the user and secure for the sender with error at most ε in both cases. The following holds:

Theorem 3.6. *If $H_{\min}(W) \geq 1$, then the identification scheme Q-ID (with suitable choice of parameters) is ε-secure against impersonation attacks for any unbounded user and for any server with quantum memory bound q, where*

$$\varepsilon = 2^{-\frac{1}{3}((\frac{1}{4} - \lambda)n\mu - 3\log(m) - q - 2)} + 2^{-(\sigma(\lambda)n\mu - \log(m) - 4)}$$

for an arbitrary $0 < \lambda < \frac{1}{4}$*, and where* $\mu = h^{-1}(1 - \log(m)/n)$*, and* h^{-1} *is the inverse function of the binary entropy function:* $h(p) := -p \cdot \log(p) - (1-p) \cdot \log(1-p)$ *restricted to* $0 < p \leq \frac{1}{2}$*. In particular, if* $\log(m)$ *is sublinear in* n*, then* ε *is negligible in* $n - 8q$*.*

Proof. We choose $\ell = \frac{1}{3}\big((\frac{1}{4} - \lambda)d + 3\log(m) - q - 1\big)$. Then user security holds except with an error $\varepsilon = 2^{-\frac{1}{3}((\frac{1}{4}-\lambda)d - 3\log(m) - q - 1)} + 2^{-(\sigma(\lambda)d - 2\ln(m) - 3)}$, and server security holds except with an error $m^2/2^\ell = 2^{-\frac{1}{3}((\frac{1}{4}-\lambda)d - 3\log(m) - q - 1)}$. Using a code $\mathfrak{c}$, which asymptotically meets the Gilbert-Varshamov bound [Tho83], d may be chosen arbitrarily close to $n \cdot h^{-1}\big(1 - \log(m)/n\big)$. In particular, we can ensure that d differs from this value by at most 1. Inserting $d = n \cdot h^{-1}\big(1 - \log(m)/n\big) - 1$ in the expression for user security yields the theorem. □

3.4 Mutual Identification

In order to obtain *mutual* identification, where also the server identifies himself towards the user, one could of course simply run *Q-ID* in both directions: say, first U identifies himself to S, and then S identifies himself to U (by exchanging their roles in *Q-ID*). However, this scheme allows the dishonest server to exclude *two* possible keys $w \in \mathcal{W}$ per invocation, and it requires to also assume the *user's* quantum memory to be bounded, and has doubled complexity.

We briefly sketch an approach that circumvents these drawbacks of the trivial solution: In the original *Q-ID* scheme, instead of announcing $z = f(x|_{I_w}) \oplus g(w)$, U announces a *noisy version* $\tilde{z}$, obtained from z by flipping each bit of z independently with some small probability; this still allows S to verify if U knows w by testing if $\tilde{z}$ is "close" to z', and S has then to prove knowledge of w by announcing to U the positions where U flipped the bits.

Security against a dishonest user still holds (with a slightly larger error probability) since the uniformity of the S_j's, as defined in the proof, also guarantees that the S_j's are pair-wise "far apart" so that W' is still uniquely determined by $\tilde{Z}$. And security against a dishonest server follows from the fact that if $W' \neq W$ then Z is (essentially) uniformly distributed and thus given its noisy version $\tilde{Z}$ the server can at best guess the positions of the bit-flips, which are independent of W.

3.5 An Error-tolerant Scheme

We now consider an imperfect quantum channel with "error rate" ϕ. The scheme *Q-ID* is sensitive to such errors in that they cause $x|_{I_w}$ and $x'|_{I_w}$ to be different and thus an honest server S is likely to reject an honest user U. This problem can be overcome by means of error-correcting techniques: U chooses a linear error-correcting code that allows to correct a ϕ-fraction of errors, and then in step 2, in addition to θ and f, U sends a description of the code and the syndrome s of $x|_{I_w}$ to S; this additional information allows S to recover $x|_{I_w}$ from its noisy version $x'|_{I_w}$ by standard techniques. However, this technique introduces a new problem: the syndrome s of $x|_{I_w}$ may give information on w to a dishonest server. Hence, to circumvent this problem, the code chosen by U must have the additional property that for a dishonest user, who has high min-entropy on $x|_{I_w}$, the syndrome s is (close to) independent of w.

This problem has been addressed and solved in the classical setting by Dodis and Smith [DS05], and subsequently in the quantum setting in [FS08b]. Dodis and Smith present a family of efficiently decodable linear codes allowing to correct a constant fraction of errors, and where the syndrome of a string is close to uniform if the string has enough min-entropy and the code is chosen at

random from the family. Specifically, Lemma 5 of [DS05] guarantees that for every $0 < \lambda < 1$ and for an infinite number of n''s there exists a δ-*biased* (as defined in [DS05]) family $\mathcal{C} = \{C_j\}_{j \in \mathcal{J}}$ of $[n', k', d']_2$-codes with $\delta < 2^{-\lambda n'/2}$, and which allows to efficiently correct a constant fraction of errors. Furthermore, Theorem 3.2 of [FS08b] (which generalizes Lemma 4 in [DS05] to the quantum setting) guarantees that if a string Y has t bits of min-entropy[8] then for a randomly chosen code $C_j \in \mathcal{C}$, the syndrome of Y is close to random and independent of j and any q-qubit state that may depend on Y, where the closeness is given by $\delta \cdot 2^{(n'+q-t)/2}$. In our application, $Y = X_W$, $n' \approx n/2$ and $t \approx d/4 - \log(m) - \ell$, where the additional loss of ℓ bits of entropy comes from learning the ℓ-bit string z. Choosing $\lambda = 1 - \frac{t}{2n'}$ gives an ensemble of code families that allow to correct a linear number of errors and the syndrome is ε-close to uniform given the quantum state, where $\varepsilon \leq 2^{-n'/2+t/4} \cdot 2^{(n'+q-t)/2} = 2^{-(t-2q)/4}$, which is exponentially small provided that there is a linear gap between t and $2q$. Thus, the syndrome gives essentially no additional information. The error rate ϕ that can be tolerated this way depends in a rather complicated way on λ, but choosing λ larger, for instance $\lambda = 1 - \frac{t+\nu q}{2n'}$ for a constant $\nu > 0$, allows to tolerate a higher error rate but requires q to be a smaller (but still constant) fraction of t.

Another imperfection has to be taken into account in current implementations of the quantum channel: imperfect sources. An imperfect source transmits more than one qubit in the same state with probability η independently each time a new transmission takes place. To deal with imperfect sources, we freely give away (x_i, θ_i) to the adversary when a multi-qubit transmission occurs in position i. It is not difficult to see that parameter ε in Proposition 3.4 then changes in that d is replaced by $(1 - \eta)d$.

It follows that a quantum channel with error-rate ϕ and multi-pulse rate η, called the (ϕ, η)-weak quantum model in [DFSS05], can be tolerated for some small enough (but constant) ϕ and η.

4 Defeating Man-in-the-Middle Attacks

4.1 The Approach

In the previous section, we "only" proved security against impersonation attacks, but we did not consider a man-in-the-middle attack, where the attacker sits between an honest user and an honest server and controls their (quantum and classical) communication. And indeed, *Q-ID* is highly insecure against such an attack: the attacker may measure the first qubit in, say, basis $+$, and then forward the collapsed qubit (together with the remaining untouched ones) and observe if S accepts the session. If not, then the attacker knows that he introduced an error and hence that the first qubit must have been encoded and measured using the $\times$-basis, which gives him one bit of information on the key w. The error-tolerant scheme seems to prevent this particular attack, but it is by no means clear that it is secure against *any* man-in-the-middle attack.

To defeat a man-in-the-middle attack that tampers with the quantum communication, we perform a check of correctness on a random subset. The check allows to detect if the attacker tampers too much with the quantum communication, and the scheme can be aborted before sensitive information is leaked to the attacker. In order to protect the classical communication, one might use a standard information-theoretic authentication code. However, the key for such a code can only be securely used a limited number of times. A similar problem occurs in QKD: even though a

[8] [FS08b] does not consider *smooth* min-entropy, but it is not too hard to see that their results also hold for the smooth version.

successful QKD execution produces fresh key material that can be used in the next execution, the attacker can have the parties run out of authentication keys by repeatedly enforcing the executions to fail. In order to overcome this problem, we will use some special authentication scheme allowing to re-use the key under certain circumstances, as discussed in Sect. 4.3.

4.2 The Setting

Similar to before, we assume that the user U and the server S share a not necessarily uniform, low-entropy key w. In order to handle the stronger security requirements of this section, we have to assume that U and S in addition share a uniform high-entropy key k. We require that a man-in-the-middle attacker can do no better that making a guess w' at w, and if his guess is incorrect then he learns no more information on w besides that his guess is wrong, and essentially no information on k. More formally:

Definition 4.1. *We say that an identification protocol is secure against man-in-the-middle attacks by* E *with error* ε *if, whenever the initial state of* E *is independent of the keys* W *and* K*, there exists* W'*, independent of* W*, such that the common state* ρ_{KWE} *after the execution of the protocol satisfies*

$$\rho_{KWW'E|W'\neq W} \approx_\varepsilon \rho_K \otimes \rho_{W\leftrightarrow W'\leftrightarrow E|W'\neq W}.$$

Furthermore, we require security against impersonation attacks, as defined in the previous section, *even if the dishonest party knows* k. It follows that k can for instance be stored on a smart card, and security is still guaranteed even if the smart card gets stolen, assuming that the theft is noticed and the corresponding party does/can not execute the scheme anymore. We would also like to stress that by our security notion, not only w but also k may be safely reused, even if the scheme was under attack.

4.3 An Additional Tool: Extractor MACs

An important tool used in this section is an authentication scheme, i.e., a Message Authentication Code (MAC), that also acts as an extractor, meaning that if there is high min-entropy in the message, then the key-tag pair cannot be distinguished from the key and a random tag. Such a MAC, introduced in [DKRS06], is called an extractor MAC, EXTR-MAC for short. For instance $MAC^*_{\alpha,\beta}(x) = [\alpha x] + \beta$, where $\alpha, x \in GF(2^n)$, $\beta \in GF(2^\ell)$ and $[\,.\,]$, denotes truncation to the ℓ first bits, is an EXTR-MAC: impersonation and substitution probability are $1/2^\ell$, and, for an arbitrary message X and "side information" Z, a random key $K = (A, B)$ and the corresponding tag $T = [A \cdot X] + B$, the tuple (T, K, Z) is $\big(\frac{1}{2} \cdot 2^{-\frac{1}{2}(H^\varepsilon_{\min}(X|Z)-\ell)} + 2\varepsilon\big)$-close to (U, K, Z), where U is the uniform distribution, respectively, ρ_{TKZE} is $\big(\frac{1}{2} \cdot 2^{-\frac{1}{2}(H^\varepsilon_{\min}(X|Z)-q-\ell)} + 2\varepsilon\big)$-close to $\frac{1}{2^\ell}\mathbb{I}\otimes\rho_{KZE} = \frac{1}{2^\ell}\mathbb{I} \otimes \rho_K \otimes \rho_{ZE}$ if we allow a q-qubit state E that may depend only on X and Z. A useful feature of an EXTR-MAC is that if an adversary gets to see the tag of a message on which he has high min-entropy, then the key for the MAC can be safely re-used (sequentially). Indeed, closeness of the real state, ρ_{TKE}, to the ideal state, $\frac{1}{2^\ell}\mathbb{I} \otimes \rho_{KE} = \frac{1}{2^\ell}\mathbb{I} \otimes \rho_K \otimes \rho_E$, means that no matter how the state evolves, the real state behaves like the ideal one (except with small probability), but of course in the ideal state, K is still "fresh" and can be reused.

4.4 The Scheme

As for Q-ID, let $\mathfrak{c} : \mathcal{W} \to \{+,\times\}^n$ be the encoding function of a binary code of length n with $m = |\mathcal{W}|$ codewords and minimal distance d, and for parameter ℓ, let $\mathcal{F}$ and $\mathcal{G}$ be strongly universal-2 classes of hash functions from $\{0,1\}^n$ to $\{0,1\}^\ell$ and $\mathcal{W}$ to $\{0,1\}^\ell$, respectively. Also, let $MAC^* : \mathcal{K} \times \mathcal{M} \to \{0,1\}^\ell$ be an EXTR-MAC with an arbitrary key space $\mathcal{K}$, a message space $\mathcal{M}$ that will become clear later, and an error probability $2^{-\ell}$. Furthermore, let $\{syn_j\}_{j\in\mathcal{J}}$ be the family of syndrome functions[9] corresponding to a family $\mathcal{C} = \{C_j\}_{j\in\mathcal{J}}$ of linear error correcting codes of size $n' = n/2$, as discussed in Section 3.5: any C_j allows to efficiently correct a δ-fraction of errors for some constant $\delta > 0$, and for a random $j \in \mathcal{J}$, the syndrome of a string with $t = (\frac{1}{4} - \lambda)d - \log(m) - 3\ell$ bits of min-entropy is $2^{-(t-2q)/4}$-close to uniform (given j and any q-qubit state) for some $\lambda > 0$.

Recall, by the set-up assumption, the user U and the server S share a password $w \in \mathcal{W}$ as well as a uniform high-entropy key, which we define to be a random authentication key $k \in \mathcal{K}$. The resulting scheme Q-ID$^+$ is given in the box below.

Q-ID$^+$:

1. U picks $x \in_R \{0,1\}^n$ and $\theta \in_R \{+,\times\}^n$, and sends the n-qubit state $|x\rangle_\theta$ to S. Write $I_w := \{i : \theta_i = \mathfrak{c}(w)_i\}$.
2. S picks a random subset $T \subset \{1,\dots,n\}$ of size ℓ, it computes $c = \mathfrak{c}(w)$, replaces every c_i with $i \in T$ by $c_i \in_R \{+,\times\}$ and measures $|x\rangle_\theta$ in basis c. Let x' be the outcome, and let $test' := x'|_T$.
3. U sends θ, $j \in_R \mathcal{J}$, $s := syn_j(x|_{I_w})$, and $f \in_R \mathcal{F}$ to S.
4. S picks $g \in \mathcal{G}$, and sends T and g to U.
5. U sends $test := x|_T$, $z := f(x|_{I_w}) \oplus g(w)$ and $tag^* := MAC^*_k(\theta, j, s, f, g, T, test, z, x|_{I_w})$ to S.
6. S recovers $x|_{I_w}$ from $x'|_{I_w}$ with the help of $test$ and s, and it accepts if and only if (1) tag^* verifies correctly, (2) $test$ coincides with $test'$ wherever the bases coincide, and (3) $z = f(x|_{I_w}) \oplus g(w)$.

Proposition 4.2 (Security against man-in-the-middle). *Assume that the quantum memory of* E *is of size at most q qubits at step 3 of* Q-ID$^+$*. Then* Q-ID$^+$ *is secure against man-in-the-middle attacks by* E *with error ε, where*

$$\varepsilon = negl\big((\tfrac{1}{4} - \lambda)d - \log(m) - 2q - 3\ell\big) + negl\big(\sigma(\lambda)d - \log(m)\big) + negl(\ell)$$

for an arbitrary $0 < \lambda < \frac{1}{4}$.

Proof. We use capital letters (W, Θ, etc.) for the values (w, θ, etc.) occurring in the scheme whenever we view them as random variables, and we write X_W and X'_W for the random variables taking values $x|_{I_w}$ and $x'|_{I_w}$, respectively. To simplify the argument, we neglect error probabilities that are of order ε, as well as linear fractions that can be chosen arbitrarily small. We merely give indication of a small error by (sometimes) using the word "essentially".

First note that due to the security of the MAC and its key, if the attacker substitutes $\theta, j, s, f, g, T, test$ or z, or if S recovers an incorrect string as $x|_{I_w}$, then S will reject at the end of the protocol. We can define W' (independent of W) as in the proof of Proposition 3.4 such that if $W \neq W'$

[9] We agree on the following convention: for a bit string y of arbitrary length, $syn_j(y)$ is to be understood as $syn_j(y0\cdots0)$ with enough padded zeros if its bit length is smaller than n', and as $\big(syn_j(y'), y''\big)$, where y' consist of the first n' and y'' of the remaining bits of y, if its bit length is bigger than n'.

then X_W has essentially $d/4 - \log(m)$ bits of smooth min-entropy, given W, W' and Θ. Furthermore, given $TAG^*, F(X_W), TEST$ (as well as K, F, T, W, W' and Θ), X_W has still essentially $t = d/4 - \log(m) - 3\ell$ bits of smooth min-entropy, if $W \neq W'$. By the property of the code family $\mathcal{C}$, it follows that if $t > 2q$ with a linear gap then the syndrome $S = syn_J(X_W)$ is essentially random and independent of $J, TAG^*, F(X_W), TEST, K, F, T, W, W', \Theta$ and E, conditioned on $W \neq W'$. Furthermore, it follows from the privacy-amplifying property of MAC^* and of f that if $d/4 - \log(m) - 2\ell > q$ with a linear gap, then the set of values $(TAG^*, F(X_W))$ is essentially random and independent of $K, F, TEST, T, W, W', \Theta$ and E, conditioned on $W \neq W'$. Finally, K is independent of the rest, and E is independent of $K, F, TEST, T, W, \Theta$. It follows that $\rho_{KWW'E|W'\neq W} \approx \rho_K \otimes \rho_{W\leftrightarrow W'\leftrightarrow E|W'\neq W}$, *before* he learns S's decision to accept or reject.

It remains to argue that S's decision does not give any additional information on W. We will make a case distinction, which does not depend on w, and we will show for both cases that S's decision to accept or reject is independent of w, which proves the claim. But first, we need the following observation. Recall that outside of the test set T, S measured in the bases dictated by w, but within T in random bases. Let I'_w be the subset of positions $i \in I_w$ with $c_i = \mathfrak{c}(w)_i$ (and thus also $= \theta_i$), and let $T' = T \cap I'_w$. In other words, we remove the positions where S measured in the "wrong" basis. The size of T' is essentially $\ell/4$, and given its size, it is a random subset of I'_w of size $|T'|$. It follows from the theory of random sampling that $\nu\big(x|_{I'_w}, x'|_{I'_w}\big)$ essentially equals $\nu\big(x|_{T'}, x'|_{T'}\big)$ (except with probability negligible in the size of T'), where $\nu(\cdot,\cdot)$ denotes the fraction of errors between the two input strings. Furthermore, since the set $V = \{i \in T : \theta_i = c_i\}$ of positions where U and S compare x and x' is a superset of T' of essentially twice the size, $\nu\big(x|_V, x'|_V\big)$ is essentially lower bounded by $\frac{1}{2}\nu\big(x|_{T'}, x'|_{T'}\big)$. Putting things together, we get that $\nu\big(x|_{I'_w}, x'|_{I'_w}\big)$ is essentially upper bounded by $2\,\nu\big(x|_V, x'|_V\big)$. Also note that $\nu\big(x|_V, x'|_V\big)$ does not depend on w. We can now do the case distinction: *Case 1:* If $\nu\big(x|_V, x'|_V\big) \leq \frac{\delta}{2}$ (minus an arbitrarily small value), then $x|_{I'_w}$ and $x'|_{I'_w}$ differ in at most a δ-fraction of their positions, and thus S correctly recovers $x|_{I_w}$ (using $test = x|_T$ to get $x|_{I_w \setminus I'_w}$ and using s to correct the rest), no matter what w is, and it follows that S's decision only depends on the attacker's behavior, but not on w. *Case 2:* Otherwise, S is guaranteed to get the correct $test = x|_T$ (or else rejects) and thus rejects as $test$ and $test'$, restricted to V, differ in more than a $\frac{\delta}{2}$-fraction of their positions. Hence, S always rejects in case 2. □

For a dishonest user or server who knows k (but not w), breaking Q-ID$^+$ is equivalent to breaking Q-ID, up to a change in the parameters. Doing the maths on the parameters similarly to the proof of Theorem 3.6 (namely, choosing $\ell = \frac{1}{4}\big((\frac{1}{4} - \lambda)d + \log(m) - 2q\big)$ whence $\varepsilon = negl\big((\frac{1}{4} - \lambda)d - 7\log(m) - 2q\big)$), it then follows:

Theorem 4.3. *If $H_{\min}(W) \geq 1$, then the identification scheme Q-ID$^+$ is ε-secure against a man-in-the-middle attacker with quantum memory bound q, and, even with a leaked k, Q-ID$^+$ is ε-secure against impersonation attacks for any unbounded user and for any server with quantum memory bound q, where*

$$\varepsilon = negl\big((\tfrac{1}{4} - \lambda)\mu n - 7\log(m) - 2q\big) + negl\big(\sigma(\lambda)\mu n - \log(m)\big)$$

for $\mu = h^{-1}(1 - \log(m)/n)$ and an arbitrary $0 < \lambda < \frac{1}{4}$. In particular, if $\log(m)$ is sublinear in n, ε is negligible in $n - 16q$.

It is easy to see that Q-ID$^+$ can tolerate a noisy quantum communication up to any error rate $\phi < \delta$. Similar to the discussion in Section 3.5, tolerating a higher error rate requires the bound on

the adversary's quantum memory to be smaller but still linear in the number of qubits transmitted. Imperfect sources can also be addressed in a similar way as for Q-ID. It follows that Q-ID$^+$ can also be shown secure in the (ϕ, η)-weak quantum model provided ϕ and η are small enough constants.

5 Application to QKD

As already pointed out in Section 4.1, current QKD schemes have the shortcoming that if there is no classical channel available that is authenticated by physical means, and thus messages need to be authenticated by an information-theoretic authentication scheme, an attacker can force the parties to run out of authentication keys simply by making an execution (or several executions if the parties share more key material) fail. Even worse, even if there is no attacker, but some execution(s) of the QKD scheme fails due to a technical problem, parties could still run out of authentication keys because it may not be possible to distinguish between an active attack and a technical failure. This shortcoming could make the technology impractical in situations where denial of service attacks or technical interruptions often occur.

The identification scheme Q-ID$^+$ from the previous section immediately gives a QKD scheme *in the bounded-quantum-storage model* that allows to re-use the authentications key(s). Actually, we can inherit the key-setting from Q-ID$^+$, where there are two keys, a human-memorizable password and a uniform, high-entropy key, where security is still guaranteed even if the latter gets stolen and the theft is noticed. In order to agree on a secret key sk, the two parties execute Q-ID$^+$, and extract sk from $x|_{I_w}$ by applying yet another strongly universal-2 function, for instance chosen by U in step 3 and authenticated together with the other information in Step 5. Here, n needs to be increased accordingly to have the additional necessary amount of entropy in $x|_{I_w}$. The analysis of Q-ID$^+$ immediately implies that if honest S accepts, then he is convinced that he shares sk with the legitimate U which knows w. In order to convince U, S can then use part of sk to one-time-pad encrypt w, and send it to U. The rest of sk is then a secure secret key, shared between U and S. In order to have a better "key rate", instead of using sk (minus the part used for the one-time-pad encryption) as secret key, one can also run a standard QKD scheme on top of Q-ID$^+$ and use sk as a one-time authentication key.

Bibliography

[ADR02] Yonatan Aumann, Yan Zong Ding, and Michael O. Rabin. Everlasting security in the bounded storage model. *IEEE Transactions on Information Theory*, 48(6):1668–1680, June 2002.

[BCS09] Harry Buhrman, Matthias Christandl, and Christian Schaffner. Impossibility of two-party secure function evaluation. in preparation, 2009.

[CCM98] C. Cachin, C. Crépeau, and J. Marcil. Oblivious transfer with a memory-bounded receiver. In *39th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 493–502, 1998.

[CW08] Claude Crépeau and Jürg Wullschleger. Statistical security conditions for two-party secure function evaluation. In *Third International Conference on Information Theoretic Security (ICITS)*, pages 86–99, 2008.

[DFL+09] Ivan B. Damgård, Serge Fehr, Carolin Lunemann, Louis Salvail, and Christian Schaffner. Improving the security of quantum protocols. `http://arxiv.org/abs/0902.3918`, 2009.

[DFR+07] Ivan B. Damgård, Serge Fehr, Renato Renner, Louis Salvail, and Christian Schaffner. A tight high-order entropic quantum uncertainty relation with applications. In *Advances in Cryptology—CRYPTO '07*, volume 4622 of *Lecture Notes in Computer Science*, pages 360–378. Springer, 2007.

[DFSS05] Ivan B. Damgård, Serge Fehr, Louis Salvail, and Christian Schaffner. Cryptography in the bounded quantum-storage model. In *46th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 449–458, 2005. Full version available at: `http://arxiv.org/abs/quant-ph/0508222v2`.

[DFSS07] Ivan B. Damgård, Serge Fehr, Louis Salvail, and Christian Schaffner. Secure identification and QKD in the bounded-quantum-storage model. In *Advances in Cryptology—CRYPTO '07*, volume 4622 of *Lecture Notes in Computer Science*, pages 342–359. Springer, 2007.

[DKRS06] Yevgeniy Dodis, Jonathan Katz, Leonid Reyzin, and Adam Smith. Robust fuzzy extractors and authenticated key agreement from close secrets. In *Advances in Cryptology—CRYPTO '06*, volume 4117 of *Lecture Notes in Computer Science*, pages 232–250. Springer, 2006.

[DM04] Stefan Dziembowski and Ueli M. Maurer. On generating the initial key in the bounded-storage model. In *Advances in Cryptology—EUROCRYPT '04*, volume 3027 of *Lecture Notes in Computer Science*, pages 126–137. Springer, 2004.

[DS05] Yevgeniy Dodis and Adam Smith. Correcting errors without leaking partial information. In *37th Annual ACM Symposium on Theory of Computing (STOC)*, pages 654–663, 2005.

[EPT03] Chip Elliott, David Pearson, and Gregory Troxel. Quantum cryptography in practice. In *SIGCOMM '03: Proceedings of the 2003 conference on Applications, technologies, architectures, and protocols for computer communications*, pages 227–238, 2003.

[FFS87] Uriel Feige, Amos Fiat, and Adi Shamir. Zero knowledge proofs of identity. In *19th Annual ACM Symposium on Theory of Computing (STOC)*, pages 210–217, 1987.

[FS86] Amos Fiat and Adi Shamir. How to prove yourself: Practical solutions to identification and signature problems. In *Advances in Cryptology—CRYPTO '86*, volume 263 of *Lecture Notes in Computer Science*, pages 186–194. Springer, 1986.

[FS08a] Serge Fehr and Christian Schaffner. Composing quantum protocols in a classical environment. `http://arxiv.org/abs/0804.1059`, 2008.

[FS08b] Serge Fehr and Christian Schaffner. Randomness extraction via delta-biased masking in the presence of a quantum attacker. In *Theory of Cryptography Conference (TCC)*, volume 4948 of *Lecture Notes in Computer Science*, pages 465–481. Springer, 2008.

[FS09] Serge Fehr and Christian Schaffner. Composing quantum protocols in a classical environment. In *Theory of Cryptography Conference (TCC)*, volume 5444 of *Lecture Notes in Computer Science*, pages 350–367. Springer, 2009.

[GL03] Rosario Gennaro and Yehuda Lindell. A framework for password-based authenticated key exchange. In *Advances in Cryptology—EUROCRYPT '03*, volume 2656 of *Lecture Notes in Computer Science*, pages 524–543. Springer, 2003.

[KOY01] Jonathan Katz, Rafail Ostrovsky, and Moti Yung. Efficient password-authenticated key exchange using human-memorable passwords. In *Advances in Cryptology—EUROCRYPT '01*, volume 2045 of *Lecture Notes in Computer Science*, pages 473–492. Springer, 2001.

[KRS08] Robert König, Renato Renner, and Christian Schaffner. The operational meaning of min- and max-entropy. `http://arxiv.org/abs/0807.1338`, 2008.

[KWW09] Robert König, Stephanie Wehner, and Jürg Wullschleger. Unconditional security from noisy quantum storage. `http://arxiv.org/abs/0906.1030`, 2009.

[Lo97] Hoi-Kwong Lo. Insecurity of quantum secure computations. *Physical Review A*, 56(2):1154–1162, 1997.

[Mau90] Ueli M. Maurer. A provably-secure strongly-randomized cipher. In *Advances in Cryptology—EUROCRYPT '90*, volume 473 of *Lecture Notes in Computer Science*, pages 361–373. Springer, 1990.

[NC00] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*. Cambridge university press, 2000.

[NPS07] Jesper Buus Nielsen, Thomas B. Pedersen, and Louis Salvail. Secure two-party quantum computation against semi-honest adversaries. In preparation, 2007.

[Ren05] Renato Renner. *Security of Quantum Key Distribution*. PhD thesis, ETH Zürich (Switzerland), September 2005. `http://arxiv.org/abs/quant-ph/0512258`.

[RK05] Renato Renner and Robert König. Universally composable privacy amplification against quantum adversaries. In *Theory of Cryptography Conference (TCC)*, volume 3378 of *Lecture Notes in Computer Science*, pages 407–425. Springer, 2005.

[STW08] Christian Schaffner, Barbara M. Terhal, and Stephanie Wehner. Robust cryptography in the noisy-quantum-storage model. `http://arxiv.org/abs/0807.1333`, to appear in *Quantum Information & Computation (QIC)*, 2008.

[Tho83] Christian Thommesen. The existence of binary linear concatenated codes with Reed-Solomon outer codes which asymptotically meet the Gilbert-Varshamov bound. *IEEE Transactions on Information Theory*, 29(6):850–853, 1983.

[WST08] Stephanie Wehner, Christian Schaffner, and Barbara M. Terhal. Cryptography from noisy storage. *Physical Review Letters*, 100(22):220502, 2008.

A Proofs

A.1 Proof of Lemma 2.1

Writing $p = P[\mathcal{E}]$ and $\bar{p} = P[\bar{\mathcal{E}}]$ we indeed get

$$\begin{aligned}
\rho_{X\leftrightarrow Y\leftrightarrow E} &= \sum_{x,y} P_{XY}(x,y)|x\rangle\langle x| \otimes |y\rangle\langle y| \otimes \rho_E^y \\
&= \sum_{x,y} \big(p \cdot P_{XY|\mathcal{E}}(x,y) + \bar{p} \cdot P_{XY|\bar{\mathcal{E}}}(x,y)\big)|x\rangle\langle x| \otimes |y\rangle\langle y| \otimes \big(p \cdot \rho_{E|\mathcal{E}}^y + \bar{p} \cdot \rho_{E|\bar{\mathcal{E}}}^y\big) \\
&= p^2 \cdot \sum_{x,y} P_{XY|\mathcal{E}}(x,y)|x\rangle\langle x| \otimes |y\rangle\langle y| \otimes \rho_{E|\mathcal{E}}^y + (1-p^2) \cdot \tau \\
&= p^2 \cdot \rho_{X\leftrightarrow Y\leftrightarrow E|\mathcal{E}} + (1-p^2) \cdot \tau
\end{aligned}$$

for some density matrix τ. If $\mathcal{E}$ is independent of X and Y, so that $P_{XY} = P_{XY|\mathcal{E}} = P_{XY|\bar{\mathcal{E}}}$, then

$$\begin{aligned}
\rho_{X\leftrightarrow Y\leftrightarrow E} &= \sum_{x,y} P_{XY}(x,y)|x\rangle\langle x| \otimes |y\rangle\langle y| \otimes \rho_E^y \\
&= \sum_{x,y} P_{XY}(x,y)|x\rangle\langle x| \otimes |y\rangle\langle y| \otimes \left(p \cdot \rho_{E|\mathcal{E}}^y + \bar{p} \cdot \rho_{E|\bar{\mathcal{E}}}^y\right) \\
&= p \cdot \sum_{x,y} P_{XY|\mathcal{E}}(x,y)|x\rangle\langle x| \otimes |y\rangle\langle y| \otimes \rho_{E|\mathcal{E}}^y + \bar{p} \cdot \sum_{x,y} P_{XY|\bar{\mathcal{E}}}(x,y)|x\rangle\langle x| \otimes |y\rangle\langle y| \otimes \rho_{E|\bar{\mathcal{E}}}^y \\
&= p \cdot \rho_{X\leftrightarrow Y\leftrightarrow E|\mathcal{E}} + \bar{p} \cdot \rho_{X\leftrightarrow Y\leftrightarrow E|\mathcal{E}} \,.
\end{aligned}$$

□

A.2 Proof of Lemma 2.2

For any pair $i \neq j$ let $\mathcal{E}_{ij}$ be an event such that $P[\mathcal{E}_{ij}] \geq 1 - \varepsilon$ and

$$\sum_z P_Z(z) \cdot \max_{x_i,x_j} P_{X_iX_j\mathcal{E}_{ij}|Z}(x_i,x_j|z) \leq 2^{-\alpha} \tag{2}$$

for all $x_i \in \mathcal{X}_i$, $x_j \in \mathcal{X}_j$ and $z \in \mathcal{Z}$. By assumption, such events exist.[10] For any $j = 1,\ldots,m-1$ define

$$L_j = \{(x_1,\ldots,x_m,z) : P_{X_1|Z}(x_1|z),\ldots,P_{X_{j-1}|Z}(x_{j-1}|z) < 2^{-\alpha/2} \wedge P_{X_j|Z}(x_j|z) \geq 2^{-\alpha/2}\}$$

Informally, L_j consists of the tuples $(x_1,\ldots,x_m,z)$, where x_j has "large" probability given z whereas all previous entries have small probabilities. We define V as follows. We let V be the index $j \in \{1,\ldots,m-1\}$ such that $(X_1,\ldots,X_m,Z) \in L_j$, and in case there is no such j we let V be m. Note that if there does exist such an j then it is unique.

We need to show that this V satisfies the claim. Fix $j \in \{1,\ldots,m\}$. Clearly, for $i < j$,

$$\begin{aligned}
\sum_z P_Z(z) \cdot \max_{x_i} P_{X_iV\mathcal{E}_{ij}|Z}(x_i,j|z) &\leq \sum_z P_Z(z) \cdot \max_{x_i} P_{X_iV|Z}(x_i,j|z) \\
&= \sum_z P_Z(z) \cdot \max_{x_i} P_{X_i|Z}(x_i|z) P_{V|X_iZ}(j|x_i,z) < 2^{-\alpha/2} \,.
\end{aligned} \tag{3}$$

Indeed, either $P_{X_i|Z}(x_i|z) < 2^{-\alpha/2}$ or $P_{V|X_iZ}(j|x_i,z) = 0$ by definition of V. Consider now $i > j$. Note that

$$\begin{aligned}
\sum_z P_Z(z) \cdot \max_{x_i} P_{X_iV\mathcal{E}_{ij}|Z}(x_i,j|z) &= \sum_z P_Z(z) \cdot \max_{x_i} \sum_{x_j} P_{X_iX_jV\mathcal{E}_{ij}|Z}(x_i,x_j,j|z) \\
&\leq 2^{\alpha/2} \sum_z P_Z(z) \cdot \max_{x_i,x_j} P_{X_iX_j\mathcal{E}_{ij}|Z}(x_i,x_j|z) \leq 2^{-\alpha/2} \,,
\end{aligned} \tag{4}$$

where the last inequality follows from the assumption (2) and the first is a consequence of the fact that the number of non-zero summands (in the sum over x_j) cannot be larger than $2^{\alpha/2}$, because

[10] In case $\varepsilon = 0$, i.e., α lower bounds the ordinary (rather then the smooth) min-entropy, the $\mathcal{E}_{ij}$ are the events "that always occur" and can be ignored from the rest of the analysis.

for any x_j with $P_{X_iX_jV\mathcal{E}_{ij}|Z}(x_i, x_j, j|z) > 0$, it also holds that $P_{X_j|Z}(x_j|z) \geq 2^{-\alpha/2}$ and the sum over all those x_j would exceed 1 if there were more than $2^{\alpha/2}$ summands. Note that per-se, $\mathcal{E}_{ij}$ is only defined in the probability space given by X_i, X_j and Z, but it can be naturally extended to the probability space given by $X_1, \ldots, X_n, Z, V$ by assuming it to be independent of anything else when given X_i, X_j, Z, so that e.g. $P_{X_iV\mathcal{E}_{ij}|Z}$ is indeed well-defined.

Consider now an independent random variable W with $H_{\min}(W) \geq 1$. By the assumptions on W it holds that $P[V \neq W] \geq \frac{1}{2}$ and $P_{X_WVWZ}(x_i, j, i, z) = P_{X_iVWZ}(x_i, j, i, z) = P_{X_iVZ}(x_i, j, z)P_W(i)$. In the probability space determined by the random variables $X_1, \ldots, X_n, V, W, Z$ and all of the events $\mathcal{E}_{ij}$, define the event $\mathcal{E}$ as $\mathcal{E} := \mathcal{E}_{WV}$, so that $P_{X_WVW\mathcal{E}|Z}(x_i, j, i|z) = P_{X_iVW\mathcal{E}_{ij}|Z}(x_i, j, i|z) = P_{X_iV\mathcal{E}_{ij}|Z}(x_i, j|z)P_W(i)$. Note that

$$P[\bar{\mathcal{E}}] = \sum_{i,j} P_{VW\bar{\mathcal{E}}_{WV}}(j, i) = \sum_{i,j} P_{V\bar{\mathcal{E}}_{ij}}(j)P_W(i) \leq \sum_{i,j} P[\bar{\mathcal{E}}_{ij}]P_W(i) \leq m\varepsilon$$

and thus $P[\bar{\mathcal{E}}|V \neq W] \leq P[\bar{\mathcal{E}}]/P[V \neq W] \leq 2m\varepsilon$. From the above, it follows that

$$\begin{aligned} p_{\text{guess}}(X_W, \mathcal{E}|VWZ, V \neq W) &= \sum_{z,i,j} \max_x P_{X_WVWZ\mathcal{E}|V\neq W}(x, j, i, z) \leq 2 \sum_{z,i\neq j} \max_x P_{X_WVWZ\mathcal{E}}(x, j, i, z) \\ &= 2 \sum_{z,i\neq j} P_Z(z) \cdot \max_x P_{X_WVW\mathcal{E}|Z}(x, j, i|z) = 2 \sum_{z,i\neq j} P_Z(z) \cdot \max_{x_i} P_{X_iV\mathcal{E}_{ij}|Z}(x_i, j|z) \cdot P_W(i) \\ &= 2 \sum_i P_W(i) \sum_{j\neq i} \sum_z P_Z(z) \cdot \max_{x_i} P_{X_iV\mathcal{E}_{ij}|Z}(x_i, j|z) \leq 2m \cdot 2^{-\alpha/2}\,, \end{aligned}$$

where we used (3) and (4) in the last inequality. The claim now follows by definition of $H_{\min}$. □